Marco Minozzo e Annamaria Guolo

Laurea in Economia del Commercio Internazionale
Laurea in Economia e Amministrazione delle Imprese
Università degli Studi di Verona (sede di Vicenza)

Statistica Inferenziale
Esercitazione 1. Stima puntuale

Esercizio A. Una ditta ha bisogno di trasportare giornalmente merci nella tratta Verona-Roma. Per fare questo noleggia uno spazio merci su ogni volo aereo, operato da una seconda compagnia. Il peso delle merci da trasportare varia da un volo all'altro a seconda delle esigenze. Sia T la variabile aleatoria che descrive il peso totale in kilogrammi delle merci da spedire in un volo. Si assuma che T sia distribuita come una distribuzione normale di valore atteso pari a 150 kg e varianza pari a 25. Il contratto con la compagnia aerea prevede che la ditta spedisca fino a 150 kg di peso per ogni volo. Nel caso in cui il peso totale delle merci spedite superi questo valore, la ditta deve pagare una multa proporzionale al quadrato della differenza fra il peso effettivo e il peso previsto di 150 kg. Nel caso in cui invece la ditta non utilizzi a pieno lo spazio noleggiato, l'inefficienza allocativa genera un costo per la ditta proporzionale al quadrato della differenza fra il peso effettivo e il peso previsto di 150 kg. In entrambe le situazioni il costo aggiuntivo in euro per ogni volo aumenta in modo proporzionale al quadrato della differenza fra il peso effettivo e il peso previsto di 150 kg, secondo la seguente legge: $C = 6,00 \cdot (T - \mathrm{E}(T))^2$.

a) Si determini la probabilità che nel prossimo volo il costo aggiuntivo C sia superiore a 576,15 euro (suggerimento: si noti che $C/(6,00 \cdot \mathrm{Var}(T))$ ha una distribuzione chi-quadrato con 1 g.d.l. (grado di libertà)).

b) Si determini la probabilità che il costo aggiuntivo totale dei prossimi 12 voli sia non superiore a 3.500,55, assumendo che il costo aggiuntivo di un volo sia indipendente da quello degli altri voli.

Sapendo che la ditta guadagna 10,00 euro per ogni kg spedito, sia G la variabile aleatoria che descrive il guadagno derivante da un volo e G_T la variabile aleatoria che descrive il guadagno derivante da sei voli.

c) Si determini la probabilità che la differenza in valore assoluto fra i guadagni in due voli successivi sia maggiore di 120,00 euro.

d) Si determini la probabilità che il guadagno G_T derivante dai prossimi sei voli sia superiore a 9.500,00 euro.

La ditta ha inoltre stimato che, per ogni volo, il costo derivante dalle spese di spedizione è una v. a. normale indipendente dal peso della merce spedita. Sia U la v. a. che descrive questo costo. Si assuma che $\mathrm{E}(U) = 120,00$ e $\mathrm{Var}(U) = 36,00^2$. Sia U_T il costo delle spese di spedizione per 6 voli e C_T il costo aggiuntivo in 6 voli.

e) Si determini la probabilità che il rapporto $R = (U_T - \mathrm{E}(U_T))/\sqrt{C_T}$ sia superiore a 13,99.

f) Si determini il valore r tale che $P(R \leq r) = 0,90$.

Esercizio B. Le Ferrovie dello Stato vogliono verificare se i recenti aumenti del prezzo della benzina hanno portato ad un incremento del numero di passeggeri nei giorni feriali del treno Verona-Milano in partenza alle 14.00. Sia X_A la v. a. che descrive il numero di passeggeri su questo treno nei giorni feriali precedenti all'aumento del prezzo e X_B la v. a. che descrive il numero dei passeggeri di questo treno nei giorni feriali successi all'aumento. Si supponga che entrambe le v. a. siano distribuite normalmente con valore atteso e varianza, rispettivamente, μ_A, σ_A^2 e μ_B, σ_B^2. L'interesse è nella differenza $\mu_B - \mu_A$. Si vuole procedere all'estrazione di un campione casuale $X_{A1}, \ldots, X_{A30}$ di 30 giorni feriali precedenti all'aumento, e di un campione casuale $X_{B1}, \ldots, X_{B20}$ di 20 giorni feriali successivi all'aumento. Sia $\bar{X}_A$ la v. a. che descrive il numero medio di passeggeri per treno del primo campione e $\bar{X}_B$ la v. a. che descrive il numero medio di passeggeri per treno del secondo campione. Siano inoltre S_A^2 e S_B^2 le v. a. che descrivono le rispettive varianze campionarie corrette. Infine si considerino $D_A^2 = \sum_{i=1}^{30}(X_{Ai} - \mu_A)^2$ e $D_B^2 = \sum_{i=1}^{20}(X_{Bi} - \mu_B)^2$.

a) Si determini la probabilità $P\Big(|\bar{X}_A - \mu_A| \geq 0,5 \cdot \sqrt{\sigma_A^2}\Big)$.

b) Si determini il valore y tale che $P\left(\frac{|\bar{X}_A - \mu_A|}{\sqrt{S_A^2}} \leq y\right) = 0,95$.

c) Si determini il valore z tale che $P\left(\frac{|\bar{X}_A - \mu_A|}{\sqrt{D_A^2}} \leq z\right) = 0,95$.

Nell'ipotesi che $\mu_A = 60$, $\mu_B = 75$ e $\sigma_A^2 = 10$, $\sigma_B^2 = 15$:

d) Determinare $P[(\bar{X}_B - \bar{X}_A) - (\mu_B - \mu_A) \leq 2]$.

e) Determinare il valore d tale che $P(S_A^2 \geq d) = 0,05$.

Esercizio C. Una società che emette carte di credito al consumo vuole valutare se il fatto che il titolare della carta abbia o meno figli influenzi la probabilità che esso sia "attivo" (ovvero la società ha un credito nei suoi confronti) un anno dopo l'apertura del rapporto. Sia p_A la probabilità che le persone senza figli restino attive un anno dopo, e p_B la probabilità che le persone con almeno un figlio restino attive un anno dopo. Si estrae un campione di 50 unità fra i titolari di carta che non hanno figli e si trova che di essi 23 sono attivi un anno dopo. Sia $\hat{p}_A = 23/50$. Si estrae inoltre un campione casuale di 50 unità fra i titolari che hanno almeno un figlio e si trova che di essi 27 sono attivi un anno dopo. Sia $\hat{p}_B = 27/50$.

a) Supponendo che $p_A = p_B$ valutare la probabilità che si abbia una differenza fra le frequenze osservate $\hat{p}_B - \hat{p}_A$ superiore a quella verificatasi fra i due campioni estratti.

Esercizio D. Sia dato un campione di 3 osservazioni dalle variabili aleatorie X_1, X_2, X_3, indipendenti e con media μ e varianza σ^2. Si considerino i seguenti due stimatori di μ: $T_1 = (x_1 + x_2 + 3x_3)/5$ e $T_2 = (2x_1 + x_3)/3$. Si valutino distorsione e varianza degli stimatori e si indichi quale dei due sia preferibile come stimatore di μ.

Esercizio E. Una società di leasing acquista macchinari per la tessitura del cotone e li dá in locazione a diverse ditte. Il valore residuo dei macchinari alla fine del primo anno è funzione del tempo di uso. Sia X_i la v. a. che descrive il numero di ore di uso del macchinario i-esimo e siano μ_X e $\sigma_X^2 = 200$ la sua media e la sua varianza, rispettivamente. Inoltre, sia R_i la perdita percentuale di valore dell'i-esimo macchinario nel primo anno e si assuma $R_i = 5 + 0,5X_i$. Si sappia infine che l'estrazione di un campione casuale di 50 macchinari ha fornito una media campionaria per le ore di uso dei macchinari pari a $\bar{x} = 110$.

a) Si dica se $\bar{R} = 5 + 0,5\bar{X}$ è uno stimatore corretto del valore atteso $\mathrm{E}(R_i)$.

b) Si determini la varianza dello stimatore $\bar{R}$.

c) Si dica se $\bar{R}$ è uno stimatore consistente di $\mathrm{E}(R_i)$.

d) Si determini il valore della numerosità campionaria necessaria affinché l'errore di stima $|\bar{X} - \mu_X|$ sia non superiore a 10 con probabilità pari a 0,95 (suggerimento: si usi il teorema del limite centrale.)

Esercizio F. Si vuole stimare l'affluenza alle urne alle prossime elezioni regionali. A tale scopo, si vuole procedere all'estrazione di un campione casuale di n elettori a cui chiedere se andranno o meno a votare. Sia N la numerosità dell'elettorato attivo e p la proporzione di elettori che si recheranno a votare. Sia inoltre $\hat{p}$ la stima della probabilità p determinata sulla base del campione che si vuole estrarre.

a) Si dica se $N \times \hat{p}$ è uno stimatore corretto dell'affluenza alle urne.

b) Si determini la numerosità campionaria n necessaria affinché l'errore di stima $\hat{p} - p$ sia in valore assoluto non superiore a 0,05 con una probabilità pari a 0,95. (Suggerimento: Si noti che la numerosità campionaria n è inversamente proporzionale alla varianza dello stimatore $\hat{p}$ la quale è funzione del parametro da stimare p. Come criterio cautelare si ponga la varianza dello stimatore $\hat{p}$ pari al valore massimo possibile, ovvero pari a $(0,5)^2/n$. Seguendo questo criterio cautelare non si rischia di sottostimare n.)

Esercizio G. Siano X_1, X_2, X_3 tre v. a. indipendenti e identicamente distribuite con media pari a μ e varianza pari a σ^2. Si consideri $T = (3X_1 + 4X_2 + 2X_3)/7 + a$ come un possibile stimatore per il parametro μ.

a) Si dica se esiste un valore di a tale che T sia uno stimatore corretto di μ.

b) Valutare la distorsione di T per $a = 10$.